Understanding the AGP Roadmap for 1999

(1X, 2X, 4X Mode, AGP Pro)

November 1998
Rev. 1.1

Intel Corporation

Table of Contents

Tables

Figures

Overview

In May 1998, Intel Corporation announced the availability of the Accelerated Graphics Port (A.G.P. or AGP) specification 2.0, an extension to the existing 1.0 specification, followed by the announcement of the AGP Pro specification in August of 1998 for the workstation market. AGP specification 2.0 introduces a new chapter on 4X-transfer mode in addition to the existing 1X and 2X transfer modes introduced by AGP 1.0 specification. In the 2X-transfer mode, double-clocked data techniques were used to transfer twice the data per each AGP clock (66 MHz) as the 1X-trnsfer mode, providing a peak bandwidth of 528 MB/s. The new 4X-transfer mode uses a quad-clocked data transfer technique to transfer four times the data per each 66 MHz clock as the 1X-transfer mode. This new 4X transfer method provides a peak AGP memory bandwidth of up to 1066 MB/s or 1.06 GB/s. For further explanation of the AGP single, double and quad data transfer and clocking, see the AGP timing in the Electrical section of this document.

The AGP Pro specification is based on the 4X mode protocol, but it introduces a new set of power and cooling requirements that meets the need of the high end graphics workstation market.

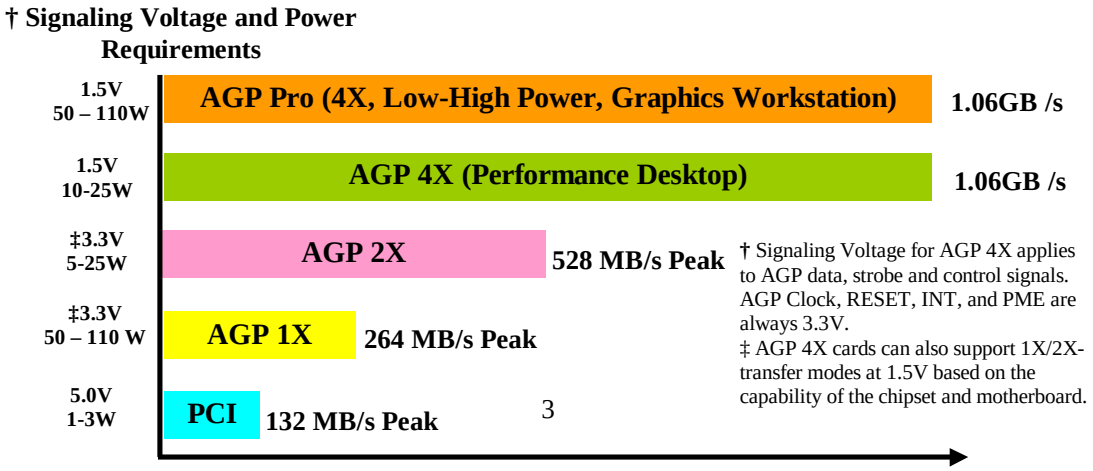
Scope

This document is intended to provide the reader sufficient information about the major differences between the 1.0 and 2.0 version of the AGP specification, in addition to the requirements set by the new AGP Pro specifications.

Evolution of the Graphics Market:

Since the introduction of the AGP specification in July 1996, the demand for higher quality 3D graphics in the desktop PC market has increased. An increasing number of 3D application and game vendors are taking advantage of the AGP memory (system memory) to store increasing amount of texture data and other attributes for realistic 3D, 2D, and video playback. AGP has moved the desktop PC graphics market into a new level of performance, previously owned by more expensive add-in graphics cards and workstations.

The success of the AGP in the desktop graphics market has motivated graphics enthusiasts to take advantage of this new bus specification, for lowering the cost of today's graphics solutions in the graphics workstation and performance PC market. The demand for this new level of performance has motivated Intel to create the AGP 2.0 (4X) specification in addition to the AGP Pro specification for the power and performance hungry workstations. Figure 1 illustrates the graphics evolution from PCI to the latest AGP 4X and AGP Pro. It should be noted that the major difference between the AGP 4X and AGP Pro is the power requirements. The protocol and the bandwidth at which they access the AGP memory are the same.



Key Differences Between AGP 1X/2X and 4X / AGP Pro

The key differences between the AGP 1.0 (1X/2X), 2.0 (4X), and AGP Pro specifications are in *Electrical, Mechanical, Thermal, Power usage*, and *Performance* characteristics. AGP cards based on the new 4X-transfer mode can be designed for high performance desktop PCs or the workstation market. The thermal requirements for the new AGP 4X cards are more stringent than for the AGP 1X/2X cards, especially for the cards in the workstation market. In order to address the power and thermal requirements of the workstation market, the AGP Pro specification offers extensions to the AGP 2.0 (4X) specification. This section will describe the major electrical and mechanical differences between the specifications.

Electrical

AGP specification 2.0 introduces a new level of signaling voltage for communication between the graphics engine and the memory controller. The new 1.5V signaling level only applies to AGP data, strobe and control signals in the 4X transfer mode. This lower voltage is required for 4X mode to ensure the integrity of the signal at such a high transfer rate. Table 1 illustrates the differences in the signaling level between 1X, 2X, and 4X modes. The next generation of Intel® AGPset will support both 1X/2X as well as 4X modes depending on the capability of the graphics engine. It is up to the motherboard manufacturer to decide what type of AGP connector to use in its design to support (1X/2X) at 3.3V, (1X/2X/4X) at 1.5V or both. (See the connector section in this document)

Signaling Level *	1X Mode	2X Mode	4X Mode
3.3 Volt	✔	✔	**NO**
1.5 Volt	✔	✔	✔

Table 1 : Allowed Transfer Mode Signaling Levels

AGP Timing

As it was explained earlier, double-clocked data techniques are used in the 2X mode and quad-clocked data techniques are used in the 4X mode to transfer twice (2X) and four times (4X) the data rate as the 1X mode per AGP 66 MHz external clock. This section explains how these techniques work in reference to the AGP external 66 MHz clock.

Single vs. Double-Clocked Data Rate

In single clock data rate (AGP 1x), the 66MHz AGP clock was used to transfer the data per rising edge of the AGP clock, allowing for 4 bytes of data to be transferred per AGP clock (4 X 66MHz = 264MB/s). In a double clocked data rate scheme (AGP 2X), STROBE signal is used to transfer the lower and upper half of the data on the rising and falling edge of the 66MHz STROB, yielding a 133MHz transfer rate (4 X 133MHz = 528MHz) . Figure 2 illustrates how Strobe signal is used in 2X mode to latch the data on its rising and falling edges per AGP clock cycle

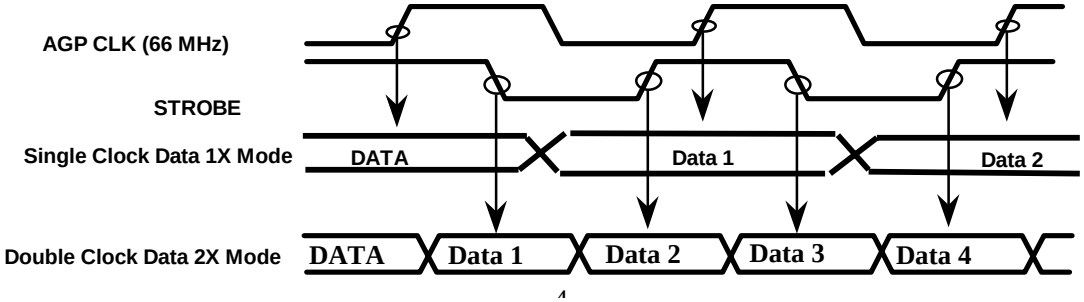


Figure 3: 1X / 2X Timing Diagram

Quad-Clock Data Rate

In the quad-clock data rate (AGP 4X) two strobes, running differentially at 133MHz are used to transfer the data on their rising and falling edges. Sine the data is latched on the falling and rising edges of a 133MHz strove, the effective transfer rate is equal to 266MHz yielding a data transfer rate of (4 X 266MHz = 1.06GB/s). Figure 3 illustrates the timing diagram used in the 4X mode. Notice that the AGP clock is still at 66 MHz.

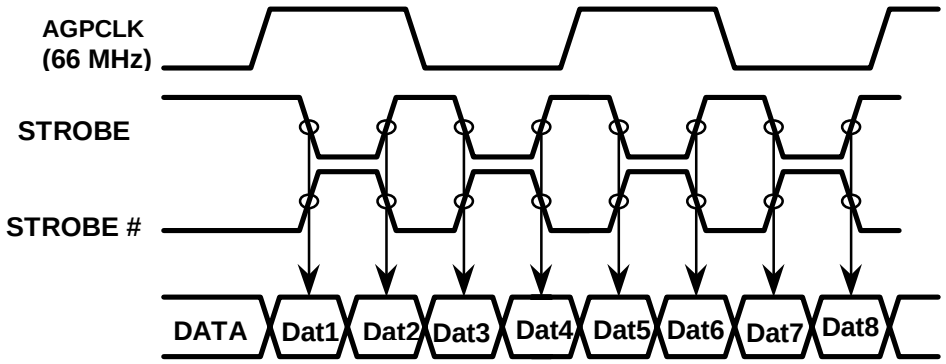


Figure 4: AGP-4X Timing Diagram

Mechanical

Due to the variation in signaling voltage, power and thermal requirements, there needs to be a mechanism to prevent a 1.5V card from being inserted into a 3.3V AGP connector and vice versa. Users also need to know what level of AGP performance their system supports. The key distinguishing factor on a motherboard will be the AGP connector. The following section explains how the four different AGP connectors in 1999 eliminate the possibility of inserting the wrong type of AGP card into a system.

AGP Connectors:

With the introduction of the AGP 4X specification and AGP Pro extensions, motherboard manufacturers can select between four different types of AGP connectors for their design, depending on their targeted market segment and level of AGP support. Figure 2 illustrates the physical differences between these four AGP connectors and their level of AGP support. The first connector to the far right is the normal AGP connector used in today's 1X /2X AGP systems. Notice the key in this connector is located in the upper section, toward the backside of the motherboard. The signaling voltage for the AGP cards that plug into this type of connector is at 3.3V. Next is the new AGP 4X connector. The keying mechanism for AGP 4X connector and cards have been moved to the lower part of the connector, closer to the front side of the motherboard, to prevent accidental insertion of a 3.3V cards into this connector and vice versa. The signaling voltage for AGP cards that plug into the AGP 4X connectors is at 1.5V.

Next is the Universal AGP connector. The Universal connector can accept either 1.5V(1X/2X/4X) or 3.3V(1X/2X) AGP cards. The idea behind the Universal connector is backward compatibility. Motherboard manufacturers that want to aim at multiple market segments and support users with older AGP cards as well as AGP 4X cards can take advantage of this new connector. The Universal connector works in conjunction

with detection and switching circuitry on the motherboard to set the right signaling voltage for AGP card that is inserted. This circuitry detects the voltage level on the TYPEDET# pin present on all AGP cards and AGP connectors to determine what type of card is inserted. Per AGP specification, the TYPEDET# pin on the AGP cards that support 1.5V signaling will be tied to ground and remain open for cards that support 3.3V signaling. During system power up (POST), the detection mechanism on the motherboard senses the TYPEDET# pin voltage levels and sets the appropriate signaling voltage. The cost and market segment advantage of motherboards that support this detection mechanism is explained in the AGP motherboard section.

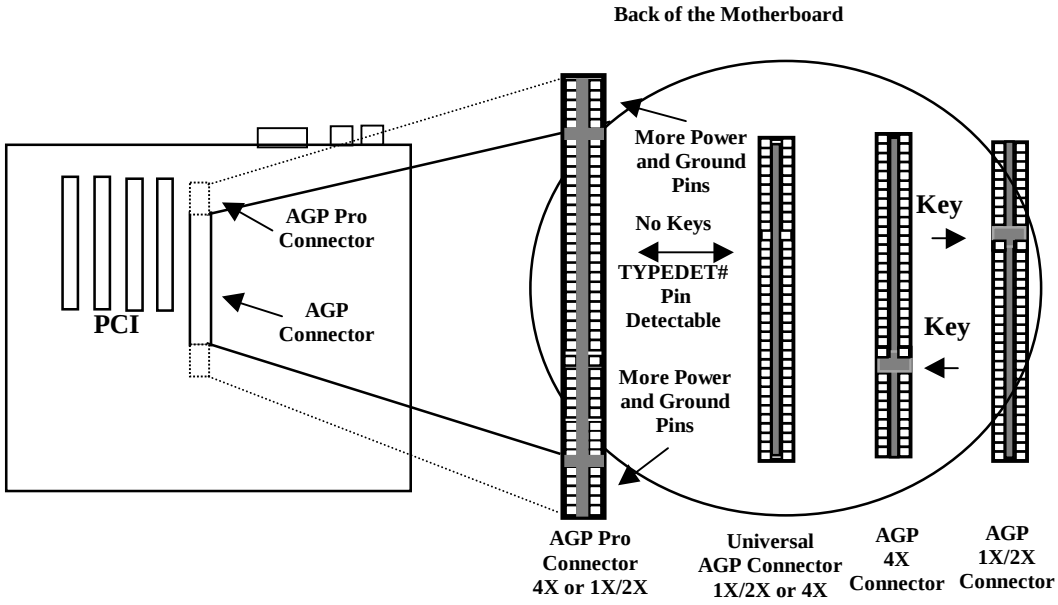


Figure 5: New Vs Current AGP Connectors

The last connector is the AGP Pro connector, designed for the workstation market. This connector has more power and ground pins, which are needed to meet the power requirement for AGP graphics cards in the workstation market. The added pins extend the length of this connector beyond the existing AGP connector as shown in Figure 2. Due to its universal design, the AGP Pro connector can accept cards that support 3.3V AGP signaling level as well as cards that support 1.5V AGP signaling. The AGP Pro cards can also be low power, using only the middle section of the connector (Universal design) or high power and use the entire connector including the additional power and ground pins. The additional power and ground pins are located at both ends, providing enough power to support advanced graphics cards with power usage in the range of 50 to 110W.

AGP Cards

The keying mechanism used in AGP connectors, works in conjunction with the cutaway section(s) in the AGP card's mating gold finger section. Figure 3 illustrates this relationship. The type of connector sets what cards can be plugged in and what AGP signaling voltage level can be applied to the card. The keying in the AGP card gold finger section can also be an identifier of what voltage the card supports. If the key as show in Figure 3 is located toward the VGA connector, it only supports 3.3V and if it is away from this connector it only supports 1.5V signaling. AGP Pro cards will have an additional register tab to the font of their gold finger section as shown in Figure 3. The high power cards will also have more power and pin fingers, which can extends the length of their gold finger section.

In the workstation market, some high performance AGP Pro graphics cards may have additional PCI cards attached to them that use the adjacent PCI connectors to increase their performance and features. AGP Pro specification requires a keep out zone of one to two adjacent PCI slots next to the AGP Pro connector for attached feature cards and cooling solutions.

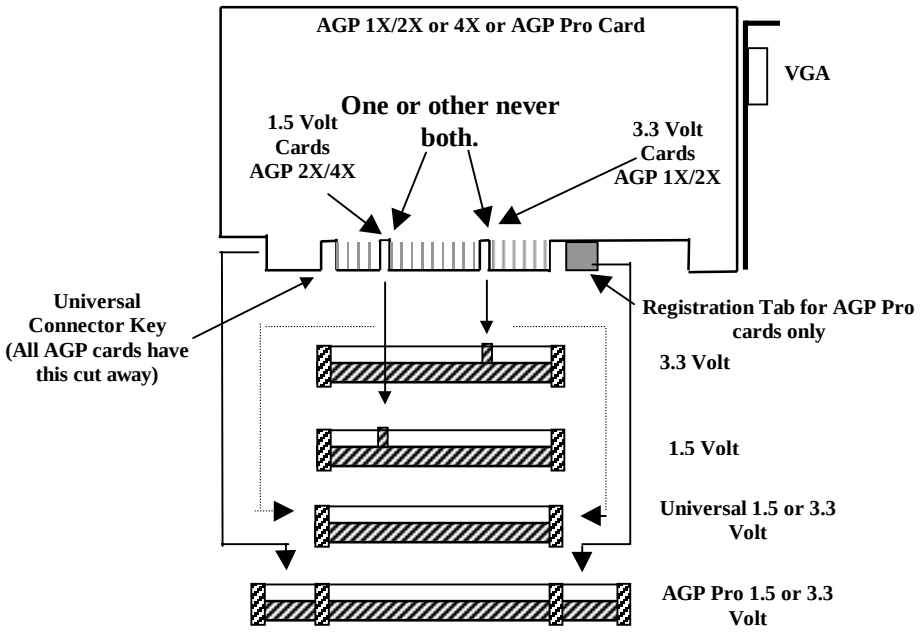
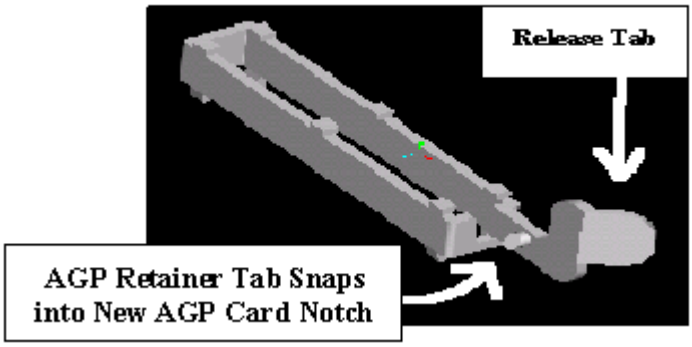


Figure 5: AGP Keying Mechanism

AGP Add-in Card Retainer

During 1997 and 1998, certain system integrators and OEMs experienced an issue with AGP cards coming out of their connectors during system shipment. There were two factors that contributed to this issue. First was the lower extraction forces applied by the AGP connector pins. Due to the high number of pins and smaller form factor than the existing PCI connector, the pins within the AGP connector had to be redesigned with smaller width, staggered position, and lower insertion and extraction forces for ease of connectivity. The second contributing factor was the actual location of the AGP connector, which was away from the back of the chassis. Due to the high speed AGP signals, the connector had to be located close to the memory controller and within the maximum length allowed by the AGP specification. The combination of these two factors contributed to the problem of AGP cards coming loose during shipment.

Intel has recognized this issue and is currently in the process of finalizing the AGP retention mechanism specification. This new specification, which should be introduced by the first part of 1999, will introduce the retainer bracket shown in Figure 4.



This new retainer bracket will snap onto the existing and upcoming AGP connectors on the motherboard (does not apply to the AGP Pro connectors). Upcoming AGP cards will have a notch at the end of their gold finger connector section, away from the VGA connector as shown in Figure 5. After inserting the new AGP cards into the connectors with retainer installed, the retainer tab shown in Figure 4 will snap into the retainer notch on the card and will lock the card in place securely. The release tab is used to unlock the card. Details about the retention mechanism will be available on (http://www.agpforum.org/) in the early part of 1999. Retainer brackets and the AGP cards that comply with this new mechanism should be available in the same time frame.

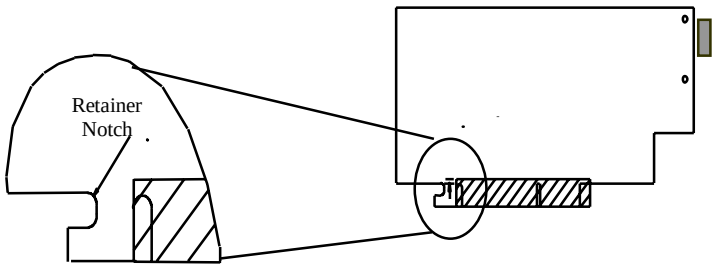


Figure 7: Retainer Notch on Upcoming AGP Cards

AGP Motherboards

As discussed in previous sections, motherboards can support 3.3V AGP cards (i.e. 1X/2X), 1.5V cards (2X/4X), or both. Additionally, in the workstation market there may be full-length, high-power AGP cards with additional PCI cards attached to them or single low power cards with no attachments. The level of AGP support on a motherboard depends on its intended market segment. For example, a motherboard that is aimed at the performance market may only support the AGP 4X at 1.5V. Another manufacturer may decide to design a motherboard for the basic PC segment and use the existing AGP connector, supporting only 3.3V cards. To expand the user market, a motherboard manufacturer may decide to use the Universal connector and provide backward compatibility to 3.3V as well as providing support for future 1.5V cards. Due to the additional TYPEDET # pin detection and voltage switching circuitry requirements on the motherboard, the Universal AGP motherboards may cost a bit more but they can reach a wider segment of the market. It should be noted that the next generation of AGP chipsets in support of AGP 4X could also provide support for 1X/2X modes for backward compatibility.

AGP Memory (System Memory)

The demand for higher system memory bandwidth by faster CPUs, I/Os and now a faster AGP has created the need to move to the next generation of the system memory technology, called Direct RDRAM or Direct Rambus. In 1999, motherboard manufacturers, aiming at the performance PC and server / workstation market, may choose to only support Direct RDRAM can provide as much as three times the memory bandwidth as today's SDRAM. Support for the current memory architecture (SDRAM) in the next generation of AGP motherboards can still be provided. As with the Universal AGP connector, there is a cost associated with supporting both types of memory due to additional circuitry, but backward compatibility creates a wider market segment. This ability to support both types of memory provides a migration path in 1999 while the production of RDRAM ramps up.

Thermal Solutions

Prior to the AGP specification, thermal management was not an issue for the desktop graphics market. With the arrival of the 2X AGP graphics, thermal solutions started to appear on some AGP graphics cards.

With the introduction of the AGP 4X specification and AGP Pro extensions, thermal solution becomes an important factor for high performance AGP graphics cards. In the workstation market, the AGP Pro compliant cards can draw up to 110W. This requires an adequate cooling system to lower the operating temperature of the graphics engine(s) on these cards.

AGP Pro extensions specify a keep out area of two PCI slots, adjacent to the AGP Pro connector (see Figure 4). In addition to providing flexibility of thermal design, the adjacent PCI slot(s) are also used for any add-in PCI card(s) that may be attached to the AGP Pro card for added features and performance. Examples of added features are real time video encoding and decoding for HDTV, added texture and geometry processing engines and etc.

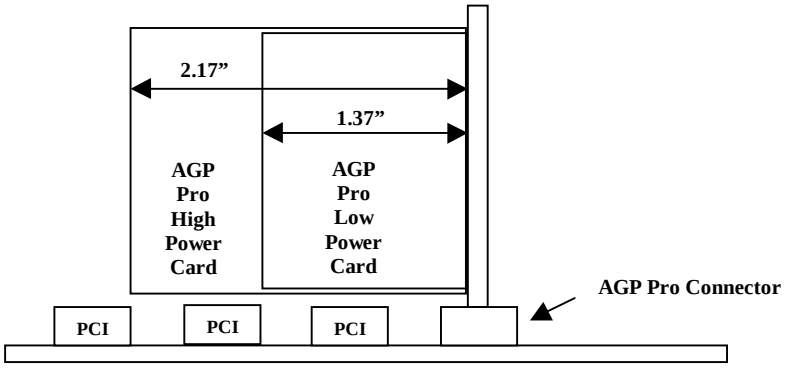


Figure 8: Orientation for Low and High Power AGP Pro Cards

Summary

Starting in the first half of 1999, system manufacturers will have to decide what level of AGP support they want to offer their customers and select their system motherboard accordingly. Will it be AGP 1X/2X at 3.3V, 2X or 4X at 1.5V, or will it have the Universal connector to accept both types? In the server market, will they offer the low power, requiring one adjacent PCI slot in the system or the high power solution, requiring two PCI slots? Once the motherboard has been selected, the connector on the motherboard becomes the governing factor for backward and forward compatibility. In the server market, the keep out area determines what type of AGP Pro card can be plugged in. AGP 4X graphics cards should start to appear in the first half of 1999. AGP card retention mechanism specification and cards should also be available in the same time frame. Table 2 summarizes the expected capabilities and specifications of AGP cards in 1999 per market segment.

AGP Connector Support on Motherboard	AGP Peak Memory Bandwidth Capability	AGP Signaling Voltage (Volt)	End User Market Segment	Average Performance	Average Power Usage (Watts)
AGP 1X/2X	264/528 MB/s	3.3	Basic PC/Mid Stream	Low to Mid	Up to 25
AGP 4X	1GB/S	1.5	Performance PC	High	Up to 25
Universal	264/528/1GB/s	3.3 or 1.5	Mid Stream / Performance	Low/Mid/High	Up to 25
AGP Pro (Low Pwr)	264/528/1GB/s	3.3 or 1.5	Low End Graphics Workstation	High to Advanced	Up to 50
AGP Pro (High Pwr)	264/528/1GB/s	3.3 or 1.5	High End Graphics Workstation	Advanced	Up to 110

Table 2: AGP Types and Specifications